

ABSTRACT OF THE DISCLOSURE

A media production system includes a textual alignment module aligning multiple speech recordings to textual lines of a script based on speech recognition results. A navigation module responds to user navigation selections respective of the textual lines of the script by communicating to the user corresponding, line-specific portions of the multiple speech recordings. An editing module responds to user associations of multiple speech recordings with textual lines by accumulating line-specific portions of the multiple speech recordings in a combination recording based on at least one of relationships of textual lines in the script to the combination recording, and temporal alignments between the multiple speech recordings and the combination recording.